

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: BILLING-MEDEL, PATRICIA A.
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
HAYDEN, MARK
KLASS, MICHAEL R.
ROBERTS-RAPP, LISA
RUSSELL, JOHN C.
STROUPE, STEPHEN D.

(ii) TITLE OF THE INVENTION: REAGENTS AND METHODS FOR THE
USEFUL FOR DETECTING DISEASES OF THE GASTROINTESTINAL
TRACT

(iii) NUMBER OF SEQUENCES: 51

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Abbott Laboratories
(B) STREET: 100 Abbott Park Road
(C) CITY: Abbott Park
(D) STATE: IL
(E) COUNTRY: USA
(F) ZIP: 60064-3500

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/828,856
(B) FILING DATE: 31-MAR-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Becker, Cheryl L.
(B) REGISTRATION NUMBER: 35,441
(C) REFERENCE/DOCKET NUMBER: 6068.US.P1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 847/935-1729
(B) TELEFAX: 847/938-2623
(C) TELEX:

10025167 - 121901

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 241 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCCAGGATA	ACTAGAGAGG	AACAATGGGG	TTATTCAAGAG	GTTTTGTTTC	CCTCTTAGTT	60
CTGTGCTCG	TGCAGAAC	AAATCTCC	TTCATTAAGC	TGAATAATAA	TGGCTTGA	120
GATATGTC	TTGTATTAAGA	TCTCTAGTCG	CCAGAAAGATG	AAAAAATATA	TGAACAAATA	180
GAGGATATGG	TGACTACAGC	TTCTACGTAC	CTGTTGAGA	CCACAGAAA	AAGATT	240
T						241

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 219 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTAGAGAGGA	ACAATGGGGT	TATTCAAGAG	TTTTGTTTC	CTCTTAGTTC	TGTGCCGT	60
GCACCAAGCTA	AAATCTCTC	TCATTAAGCT	GAATAATAAT	GGCTTGAAG	ATATTGTCAT	120
TGTTATAGAT	CCTAGTGTG	CAGAAAGATG	AAAAAATATA	GAACAAATAG	AGGATATGGT	180
GACTACAGCT	TCTACGTACC	TGTTGAGC	CACAGAAA			219

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 231 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: base_polymorphism
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "'N' represents an A or G or T or C polymorphism at this position"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTNTGTAACG	AAAAAACCCA	TAATCAAGAA	GCTCCAAGCC	TACAAACAT	AAAGTGCAAT	60
TTTACAAGTA	CATGGGAGGT	GATTAGCACT	TCTGAGGATT	TTAAAACAC	CATACCCATG	120
GTGACACCCAC	CTCCCTCCAC	TGTCCTCTCA	TTGCTGAAGA	TCAGTCAAAG	AATTGTTGTC	180
TTAGTTCTTG	ATAAGTCTGG	AAGCATGGGG	GGTAAGGACC	GCCTAAATCG	A	231

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 237 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

100025167 - 121901

TGGGGGGTAA	GGACCGCCTA	AATCGGAATG	ATCAAGCAGC	AAAACATTTC	CTGCTGCAGA	60
CTGTTGAAAAA	TGGATCCCTGG	GTCGGGATGG	TTCACTTGTG	TAGTACTGCC	ACTATTGTAA	120
ATAAGCTTAAT	CCAATAAAAA	AGCAGTGATG	AAAGAACAC	ACTCATGCCA	GGATTACTTA	180
CATATCCTCT	GGGAGGAAC	TCCATCTGCT	CTGGAAATTAA	ATATGCATT	CAGGTGA	237

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 216 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTCCATCTG	CTCTGGAAATT	AAATATGGAT	TTCAAGGTGAT	TGGAGAGGCTA	CATTCCCAC	60
TCGATGGATC	CGAAAGTACTG	CTGCTGACTG	ATGGGGAGGA	TAACACTGCA	AGTTCTTGTA	120
TTGATGTAAGT	GAAGACAAAGT	GGGGCCATTG	TTCATTATAT	TGCTTGGGA	AGAGCTGCTG	180
ATGAAAGCAT	AATAGAGATG	AGCAAGATA	CAGGAG			216

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 201 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: base_polymorphism
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "N' represents an A or G or T or C polymorphism at this position"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AATTGATAGT	ACAGTGGGAA	AGGNACGTT	CTTTCTCATC	ACATGGAACA	GTCTGCCTCC	60
CAGTATTCT	CTCTGGGATC	CCAGTGGAAC	AATAATGGAA	AATTTCACAG	TGGATGCAAC	120
TTCAAAATG	GCTTATCTCA	GTATTCAGG	AACTGCAAAG	GTGGGCACTT	GGGCATACAA	180
TCTTCAAGCC	AAAGCGAAC	C				201

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 241 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCAAATTCTT	CTGTGCCTCC	AATCACAGTG	AATGCTAAA	TGAATAAGGA	CGTAAACAGT	60
TTCCCCAGCC	CAATGATITGT	TTTACCCAGAA	ATTCTACAGG	GATATGTAAC	TGTCTTGTGA	120
GCCATGTGA	CTGTTTCTCAT	TGAATCACAG	ATTGGACATA	CAGAAGTTT	GGAACTTTTG	180
GATATGTTG	CAGCGCGTGA	TTCTTTCAAG	AATGATGGG	TCTACTCCAG	GTATTTACA	240
G						241

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 242 base pairs
- (B) TYPE: nucleic acid

10025167 - 121901

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTGCAGGCCGC	TGATTCTTTC	AAGAATGATG	GAGTCTACTC	CAGGTATTTT	ACAGCATATA	60
CAGAAAATGG	CAGATATAGC	TTAAAAGTC	GGGCTCATGG	AGGAGCAAAC	ACTGCCAGGC	120
AAAAATTACG	GCCTCCACTG	AATAGAGCCG	CGTACATACC	AGGCTGGTA	GTGAACGGGG	180
AAATTGAAAC	AAACCCGCA	AGACCTGAAN	TTGATGAGGA	TACTCAGACC	ACCTTGAGG	240
AT						242

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 233 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCGCCAAGA	CCTGAAATTG	ATGAGGATAC	TCAGACCACCC	TTGGAGGAAT	TCAGCCGAAC	60
AGCATCCGGA	GGTCGATTTC	TGGTATCACCA	AGTCCCAGC	CTTCCCTTCG	CTGACCAATA	120
CCCACCAAGT	CAAATCACAG	ACCTTGAATGC	CACAGTTCAT	GAGGATAAGA	TTATTCCTAC	180
ATGGACACGA	CCAGGAGATA	ATTITGATGT	TGGAAAAGT	CAACGTTATA	TCA	233

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 313 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: base_polymorphism
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "'N' represents an A or G or T or C polymorphism at this position"

(ix) FEATURE:
(A) NAME/KEY: base_polymorphism
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "'N' represents an A or G or T or C polymorphism at this position"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCAATACCCA	CCAAGTCAAA	TNACAGACCT	TGATGCCACA	GTINATGAGG	ATAAGATTAT	60
TCTTACATGG	ACAGCACCGAG	GAGATAATTG	TGATGTTGGA	AAAGTTCAAC	GTATATATCAT	120
AAGAATAAGT	CGAACGTTATTC	TTGATCTAAC	AGACAGTTTT	GATGATGCTC	TTCAAGTAAA	180
TACTACTGAT	CTGTCACCAA	AGGAGGCCAA	CTCCAAGGAA	AGCTTTCAT	TTAAACCAGA	240
AAATATCTCA	GAAGAAAATG	CAACCCACAT	ATTITATGCC	ATTTAAAGTA	TAGATAAAGC	300
ATTTGGCATC	AAA					313

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 242 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10025167-121901

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAGTATTCCTT GATCTAAGAG ACAGTTTGAG TGATGCTCTT CAAGTAATAA CTACTGATCT
GTCACCAAG GAGGCCAACT CCAAGGAAG CCTTGATTT AAACCGAAAA ATATCTCAGA
AGAAAATGCA ACCCACATAT TTATGCCAT TAAAAGTATA GATAAAAGCA ATTGACATC
AAAAGTATCC AACATTGCAC AAGTAACCTT GTTTATCCCT CAAGCAAATC CTGATGACAT
TG 60
120
180
240
242

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 208 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: base_polymorphism
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /note= "'N' represents an A or G or
T or C polymorphism at this position"

(ix) FEATURE:

- (A) NAME/KEY: base_polymorphism
- (B) LOCATION: 4
- (D) OTHER INFORMATION: /note= "'N' represents an A or G or
T or C polymorphism at this position"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ANANAATGCA ACCCACATAT TTATGCCAT TAAAAGTATA GATAAAAGCA ATTGACATC
AAAAGTATCC AACATGGAC AGAATCACTT GTTTATCCCT CAAGCAAATC CTGATGACAT
TGATCCACT CCTACTCCCTA CCTCTACTCC TGATAAAAAGT CATAATTCG GAGTTAATAT
TCTACGCTG GTATTGTCTG TGATTGGG 60
120
180
208

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 201 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTCTACTCC TACTCCTGAT AAAAGTCATA ATTCTGGAGT TAATATTTCT ACGCTGGTAT
TGCTCTGTAT TGGGTCTGTT GTAATGTTA ACTTTATTTT AAGTACCAACC ATTGACACT
TAACGAAGAA AAAATCTTC AAGTAGACCT AGAAGAGAGT TTAAAAAAC AAAACATGT
AAGTAAGGA TATTTCTGAA T 60
120
180
201

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 301 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: base_polymorphism
- (B) LOCATION: 111
- (D) OTHER INFORMATION: /note= "'N' represents an A or G or
T or C polymorphism at this position"

10025167 - 121901

(ix) FEATURE:
(A) NAME/KEY: base_polymorphism
(B) LOCATION: 244
(D) OTHER INFORMATION: /note= "N' represents an A or G or T or C polymorphism at this position"

(ix) FEATURE:
(A) NAME/KEY: base_polymorphism
(B) LOCATION: 284
(D) OTHER INFORMATION: /note= "N' represents an A or G or T or C polymorphism at this position"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCTGTTGTAA	TTGTAACTT	TATTTTAAGT	ACCACCATTT	GAACCTAAC	GAAGAAAAAA	60
ATCTCAAGT	AGACCTAGAA	GAGAGACAAA	AAAATCAAAG	CAATGTAAGT	NAAGGATATT	120
TCTGAATCTT	AAAAATCATC	CCTATGTGTG	TCATAAAACTC	ATAAATATAA	TTTTAAGATG	180
TCGGAAAGG	ATACTTTGAT	TAAATAAAAA	CACTCATGGA	TATGTAAGAA	CTGTCAAGAT	240
A	TAAANATTAA	TAGTTTCATT	TATTTGTAT	TTTATTGTA	AGANATAGTG	300
						301

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 229 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGAAAAAAATC	TTCAAGTAGA	CCTAGAAGAG	AGTTTTAAA	AAACAAAACA	TGTAAGTAAA	60
GGATATTTCT	GAATCTAAA	ATTCCATCCA	TGTGTGATCA	TAACACTATA	AAAATAATTT	120
TAAGATGTCG	AAAAGGATA	CITTGATTAA	ATAAAAACAC	TCATGGATAT	GTAAAACATG	180
TCAAGATTAA	AATTTAATAG	TTTCATTTAT	TTGTTATT	TTTGTGTAAG		229

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3043 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTAGAGAGGA	ACAATGGGGT	TATTCAGAGG	TTTGTTTTC	CTCTTAGTTC	TGTGCCTGCT	60
GCACCACTCA	AATACTTCCT	TCATTAAGCT	GAATAATAAT	GGCTTTGAAG	ATATTGTCAT	120
TGTTTATAGAT	CCTAGTGTGC	CAGAAGATGA	AAAATAATT	GAACAAATAG	AGGATATGGT	180
GAACAGCT	TCTACGTACC	TGTTGAAGC	CACAGAAAAA	AGATTTTTTT	TCAAAATATGT	240
ATCTATATTA	ATTCCTGAGA	ATTGGAGGA	AAATCCTCG	TACAAAAGGC	CAAAACATG	300
AAACCCATAAA	CATGCTGATG	TTATAGTTGC	ACCACTACA	CTCCCAGGTA	GAGATGAACC	360
ATACACCAAG	CAAGTTCACAG	AAATGTGGAA	GAAGGGCGAA	TCATCTTCACT	TCACCCCTG	420
CCTCTACTT	AAAAAAACAC	AAATAGATAA	TGGACCAACCA	GCCAAACTGT	TTGTCATG	480
GTGGGCTCAC	CTTCGGCTGGG	GAGTGTGTTG	TGAGTACATA	GAAGATCAGC	CTTCTTACCG	540
TGCTAAGTC	AAAAAAATCG	GTGTTCCGCA	GGTATCTCTG	TGAGAAATAG		600
AGTTTTAATAG	TGTCAGGAG	GCAGCTGTCT	TAGTAGAGCA	TGCGAAATTG	ATTCCTACAAAC	660
AAACATGTAT	GGAAAAAGATT	TCATTAATTCT	TCTCTGATAAA	GTACAAACAG	AAAAAGCATC	720
CATAATGTTT	ATGCAAGATG	TTGATTCTGT	TGTTGAATT	TGTAACGAA	AAACCCATAA	780
TCAAGAAGCT	CCAAGCCTAC	AAAACATAAA	GTGCAATT	AGAAGTACAT	GGGAGGTGAT	840
TAGCAATTCT	GAGGATTAA	AAAACACCAT	ACCCATGGTG	ACACCAACCTC	CTCCACCTGT	900

CTTCTCATTTG	CTGAAGATCA	GTCAAAGAAT	TGTGTGCTTA	GTTCTTGATA	AGTCTGGAAG	960
CATGGGGGGT	AAGGACGCC	TAATTCGAAT	GAATCAAGC	GCAAAACATT	TCCCTGTCGA	1020
GACTGTTGAA	AATGGATCCT	GGGTGGGGAT	GGTTCACCTT	GATAGTACTG	CCACATTTGT	1080
AAATAAGCTA	ATCCAAATAA	AAAGCAGTGA	TGAAAGAAC	ACACTCATGG	CAGGATTAC	1140
TACATATCCT	CTGGAGGAAAC	CTTCACATCTG	CTCTGGAATT	AAATATGCAT	TTCAAGGTGAT	1200
TGGAGAGCTA	CATTCACCAAC	TCGATGGATC	CGAAGTACTC	CTGCTGACTC	ATGGGGAGGA	1260
TAACACTGCA	AGTCTCTGTA	TTGATGGAATG	GAACAAAGATG	GGGGCCATTC	TTCATTTTAT	1320
TGCTTTGGGA	AGAGCTGCTG	ATGAGGAGCT	AATAAGAGTC	AGCAAGATAA	CAGGAGGAAG	1380
TCATTTTAT	GTTCAGATG	AAAGTCAGAA	CAATGGCTTC	ATGAGTCTT	TTGGGGCTCT	1440
TACATCAGGA	AATACTGATC	TCTCCCAAAG	GTCCCCTAG	CTCGAAAGTA	AGGGATTAAC	1500
ACTGAATAGT	ATACTCTGGA	TGAACGACAC	TGCTCAATTAA	GATAGTACAG	TGGGAAGGGA	1560
CACGTTCTT	CTCATCACAT	GGAAACAGTC	GGCTCCCGAT	ATTTCCTCTC	GGGATCCCG	1620
TGAAACATA	ATGGAAATT	TCACAGTGGA	TCGACACTTC	AAAATGGCT	ATCTCAGTAT	1680
TCCAGGAATC	GCAAGGTTG	GCACAGTGGG	ATACATCTT	CAAGCCAAAT	CGAACCCAGA	1740
AAACATAACT	ATTAACGTA	CTTCCTCGAG	AGCAATCTT	CTCTGCTTC	CAATACAGT	1800
GAATGCTAA	ATGAATTAAGG	ACGTAACAC	TTTCCCAGAT	CCAATGATTA	TTTACGCGAA	1860
AATTCTAAC	GGATGATGTC	CTGTTCTTGG	AGCCATATGTG	ACTGCTTCA	TTGATACACA	1920
GAATGGACAT	ACAGAAGTTT	TGGAACTTTT	GGATAATGGT	GCAGCGCTG	ATTCTCTTCA	1980
GAATGATGAT	GTCTACCTCA	GGTATTATAC	AGCATATACA	AAAATGGCA	GATATAGCTT	2040
AAAAGTCGG	GCTCATGGAG	GCGAACACAC	TGCGAGGCTA	AAATTACGGC	CTCCACTGAA	2100
TAGAGGCCG	TACATACCG	GCTGGGTGAT	GAACGGGGAA	ATTGAGGCAA	ACCCGCCAAG	2160
ACCTGAAATT	GATGAGGATA	CTTCAGAACAC	CTTGGAGGAT	TCACCGGAA	CAGCATCCGG	2220
AGGTGCAATT	GTGGTATCAC	AAAGTCGGAC	CTTCCCTCTG	CCTGACCAAT	ACCCCAACAG	2280
TCAAACTACA	GACCTTGTG	CCACAGTCA	TGAGGATAAG	ATTATTCTTA	CATGGACAGC	2340
ACCAGGAGAT	AATTGTGAT	TTGGGAAAGT	TCAACGTTAT	ATCATAAAGA	TAAGTGCAG	2400
TATTCTGTGAT	CTAAGAGACA	TTTTGTGATG	TCGCTCTTCA	TTAAATACTA	CTGATCTGTC	2460
ACCAAAGGAG	GCCAACTCCA	AGGAAGCTT	TGCAATTAAA	CCAGAAATA	TCTCAGAAGA	2520
AAATGCAACC	CACATATTA	TTGCGATTAA	AAGTATAGAT	AAAAGCAATT	TCGATCAAA	2580
AGTATCCAC	ATTGCAACAG	TAACCTTGT	TATCCCTCAA	GCAARTCCCG	ATGACATTGA	2640
TCCTACTCT	ACTCTACTC	CTACCTCTGA	AAAAGCAT	AATTCTGGAG	TTAATATTTC	2700
TACGCTGGTA	TTGCTGTGTA	TTGGGCTCTG	TGTAATTGTT	AATTCTATT	TAAGTACAC	2760
CATTGGAAC	TTAACGAAAG	AAAATTTCTT	CAAGTAGACAG	TAGAGAGAG	TTTTAAAAAA	2820
CAAACAAATG	TAAGTAAAGG	ATATTCTGA	ATCTTAAAT	TCATCCATG	IGTGTACATA	2880
AACTCATAA	AATAATTTTA	AGATGTCGGA	AAAGGATACT	TTGATTAAT	AAAACACTC	2940
ATGGATATG	AAAATGTC	AAGATTTAA	TTTAAATCTT	TGATTTTT	GTTATTTTAT	3000
TTGTAAGAAA	TAGTGATGAA	CAAAGATCTT	TTTCATACACT	GAT		3043

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1399 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCAAATTCTT	CTGTCCTCC	AATCACAGTC	AATGCTAAA	TGAATAAGGA	CTGAAACAGT	60
TCCCCAGGG	CAATGTTG	TTACCGAGAA	ATTCTCAACAG	GATATGTACC	TGTTCTTGG	120
GCCAT'GTGA	CTGCTTCTCAT	TGAATCACAG	AAAGGCCATCA	CAGAACTTTT	GGAACTTTTG	180
GATAATGGTC	CAGGGCGCTGA	TCTTTTCAAG	AATGATGGAG	TCTACTCCAG	GTATTTCATCA	240
GCATATACAG	AAAATGGCAG	ATATAGCTTA	AAAGTCGGGG	CTCATGGAGG	AGCAAAACACT	300
GCCAGGCTAA	AATTACGGCC	TCCACTGAAAT	AGAGCCCGCT	ACATACCGGG	CTGGCTACTG	360
AACGGGGAAA	TTGAAGCAA	CCCGCCAGA	CTCTGAAATT	ATGAGGATAC	TCAGACCCAC	420
TTGGAGGATT	TCAGCGGAAC	AGCATCCGGG	GGTGCATTTC	TGTTATCACA	AGTCCCAAGC	480
CTTCCCTTGC	CTGACCAATA	CCCCCAAGT	CAACATCACAG	ACCTTGATGC	CAACGTTCAT	540
GAGGATAAGA	TATTTCTTCA	ATGAGCAGCA	CCAGGAGATA	ATTCTGATGT	TGGAAAAGTT	600
CAACGTTATA	TCATAAGAT	AACTGCAAGT	ATTCTGATC	TTAGAGACAG	TTTTGATGAT	660
GCTCTTCAAG	TAATACTACT	TGATCTGTC	CCAAAGGAGG	CCAACTCCAA	GGAAAGCTTT	720
GCATTAAAC	CAGAAATAAT	CTCGAGAAA	AAATGCAACCC	ACATATTTAT	TGCCATTAAA	780
AGTATGATA	AAAAGCAATT	GACATAAAAA	GTATCCAACA	TTGCAACAGT	AACTTTGTTT	840
ATCCCTCAAG	CAAACTCTGA	TGACATGATG	CCCTACTCTA	CTCTTACTCC	TACTCTGTAT	900
AAAAGCTATA	ATTCTGGAGT	TAATTTCTT	ACGCTGTTAT	TGCTGTGAT	TGGGCTCTGGT	960
GTAATTGTTA	ACTTTATTTT	AAAGTACCAAC	ATTGAACTT	TAACGAAGAA	AAAATCTTC	1020

AAGTAGACCT	AGAAAGAGT	TTAAAAAAC	AAAACAATG	AGTAAGGA	TATTCGAA	1080
CTTTAAAATT	CATCCCATGT	GTGATCATAA	ACTCATAAAA	ATAATTAA	GATGTCGAA	1140
AAGGATACCT	TGATTTAATA	AAAACACTCA	TGGATATGTA	AAAACTGTCA	AGATTAATAAT	1200
TTAATAGTTT	CATTATTG	TTATTTAT	TGTAAGAAAT	AGTGATGAA	AAAGATCCTT	1260
TTCTACTG	ATACCTGGTT	GTATATTATT	TGATGCAACA	GTTTCTGAA	ATGATATTTC	1320
AAATTCATC	AAGAAATAAA	AATCATCTAT	CTGAGTAGTC	AAAATACAA	AAAGGGAGAG	1380
CAAATTAACAC	ACATTTGGA					1399

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3181 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCGGAGGATA	ACTAGAGAGG	AAACATGGGG	TTATTCAAGAG	GTTTGTGTTT	CCTCTTAGTT	60
CTGTGCTGC	TGCACCAAGTC	AAATACTTC	TTCAATTAAAGC	TGAAATAATAA	TGGCTTGGAA	120
GATATTGTC	TGTTTATAGA	TCTCTAGTGT	CCAGAAGATG	AAAATAATAT	TGACAAATA	180
GAGGAGATGG	TGACTACAGC	TCTCTAGCT	CIGTTTGAAGG	CCACAGAAAA	AAGATTTTT	240
TTCAAAATG	TATCTATATT	AAATTCGAG	AAATTGGAGG	AAAATCTCA	GTACAAAAGG	300
CCAAACATC	AAACACATTA	ACATCGTGT	GTTPATGTTG	CACCACTAC	ACTCCCCAGGT	360
AGAGATGAC	CATAACACCA	GCAGCTTACA	GAATGTTGGA	AGAAAGGCCG	ATACATTCAC	420
TTCACCCCTG	ACCTCTACT	TGAAAAAAAA	CAAATGAT	ATGGACCCG	AGGCAAATCG	480
TTTGTCCATG	AGTGGGTC	CCTCCGGTGG	GGAGGTGTTG	ATGAGTACAA	TGAGAGTCAG	540
CTCTTCATCC	GTGCTGAATG	AAAAAAAT	GAAGCACAA	GGTGTTCGCG	AGGTATCTCT	600
GGTAGAAATA	GAGTTTAA	GTGTCAGGA	GGCAGCTGTC	TTAGTAGAGC	ATGCAGAAATT	660
GATTCTACAA	CAAAACATGTA	TGGAAGAGAT	TGTCATTCTG	TTCTGTATAA	AGTACAACAA	720
GAAAAGAC	CCATAATGTT	TATGCAAAAT	ATTGATCTG	TTGTGTAATT	TTGTAAAGCAA	780
AAAACCCATT	ATCAAGAAC	TCCAAGCTTA	CAAACACAT	ATGICAAATT	TAGAAGTACA	840
TGGGAGGTGA	TTAGCAATT	TTAGGATTTT	AAAAACACCA	TACCCATGTT	GACACCACT	900
CCTCCACCTG	CTCTTCATT	GTCGATGAGT	AGTCAGAAAG	TTGTGCTCTT	AGTCTCTGT	960
AAGCTGGAA	GCATGGGGGG	TAAGGACCGC	CTAAATCGAA	TGAATCAAGG	AGCAAAACAT	1020
TTCTGCTGC	AGACTGTTGA	AAATGGATCC	TGGGTGGGGA	TTGGTTCATT	TGATAGTACT	1080
GCCACTATTG	TAATAAAGCT	AAATCCAAT	AAAGCAGTC	ATGAAAGAAA	CACACTATG	1140
GCAGGATTA	CTACATACAT	TCTGGGGAGG	ACTTCCATCT	GCTCTGGAT	TAAATATGCA	1200
TTTCAGGTTG	TTGGGAGACT	ACATTCAGT	CTCGATGGAT	CCGAAGTACT	GCTGCTGACT	1260
GATGGGAGG	ATAACACTGG	AAGTCTCTG	ATTGATGAG	TGAACAAAG	TGGGCCATT	1320
GTTCATTGTA	TTGCTTGTG	AAAGAGTCGT	GATGAGAACG	TAATAGAGAT	GAGCAAGATA	1380
ACAGGAGGAA	GTCACTTTA	TGTTCTAGAT	GAAGCTCGAA	ACATGGCTC	CATIGATGCT	1440
TTGGGGCTC	TTACATCAGG	AAATCTGAT	CTCTCCAGA	AGTCCTCTG	GCTCGGAAAGT	1500
AAGGGATTA	CACITGAATG	TAATGCTCTG	ATGAAACGACA	CTGTGATATA	TGATAGTACA	1560
GTGGGAAAGG	ACACGTTCTT	TCTCATACAA	TGGRAACGTC	TCGCTCTCCAG	TATTTCTCTC	1620
TGGGATCCCC	GTGGAACACAT	AAATGGAAAT	TTCACTAGT	ATGCAACTTC	CAAAATGGCC	1680
TATCTCTGAA	TTCCAGGAA	TGCAAAAGTG	GGCACTTGGG	CATACAATCT	TCAAGCCAAA	1740
GCGAACCCAG	AAACATTAAC	TATACAGTA	ACTTCAGGAG	CAGCAAACTT	TTCTGTGCT	1800
CCAATCACG	TGAAATGCTAA	AAATGAAATA	GACGTAACAA	GTTCCTCCAG	CCCAGTATT	1860
GTTTACGCG	AAATTCTACA	AGGATGATGA	CTCTGTTCTG	GAGCCCAATGT	GACTGCTTC	1920
ATTGAATCAC	AGAAATGGAC	TACAGAAGAT	TTGGAACCTT	TGGATAATGG	TGCAGGCGCT	1980
GATTCTTCTA	AGAATGATG	AGTCTACTGC	AGGTATTGTTA	CAGCATATAC	AGAAAATATGC	2040
AGATATAGCT	TAAAAGTCTG	GGCTCTATGA	GGAGCAAAACA	CTGCGAGGCT	AAAAATACCG	2100
CCTCCACTGTA	ATAGAGCCCG	GTACATACCA	GGCTGGGTAG	TGAACGGGGA	AATTGAGCA	2160
AACCCGCCAA	GACCTGAAAT	TGATGAGGAT	ACTCAGACCA	CCTTGGAGGA	TTTCAGCCGA	2220
ACAGCATCCG	GAGGTGCAATT	TGTGTTATCA	CAAGTCCCAA	CCCTCCCTT	GCCTGACCAA	2280
TACCCACCA	GTCAACATAC	AGACCTTGT	GCCACAGTTC	ATGAGGATAAA	GATTATCTT	2340
ACATGAGAC	CACAGGAGA	TAATTTGT	GTGAAAGAA	TTCAACGTTA	TATCATAGA	2400
ATAAGTGC	GTATTCTG	TCTAGAGAC	AGTTTGTGAT	ATGCTCTCA	AGTAAATACT	2460
ACTGATCTG	CACCAAGAA	GGCCAACTCT	AAGGAAGATC	TTGCAATTAA	ACCCAGAAAT	2520
ATCTCAGAA	AAAATGCAAC	CCACATATT	ATGCGATTA	AAAGTATGAA	TAAAAGCAAT	2580
TTGACATCAA	AAAGTCTCAA	CATTGCAACA	GTAAACTTGT	TATCTCTCA	AGCAAACTCT	2640
GATGACATG	ATCCCTACTC	TACTCTCA	CCTACTCTG	ATAAAAGTC	TAATTCTGGA	2700
GTAAATATT	CTACGCTGGT	ATTGCTGTG	ATGGGTCTG	TTGTAATTGT	TAACCTTATT	2760
TTAAGTACCA	CCATTGGAAC	CTTAAACGAAG	AAAAAAATCT	TCAAGTAGAC	CTAGAGAGA	2820

GT	TTT	AAA	AAAC	AAAT	GTA	AGT	AAAG	GAT	ATT	TCTG	AAT	CTT	AAA	TTCA	TCC	CAT	CCC	AT	2880	
TG	TG	TG	TG	TG	ATC	ACT	CATA	AA	AT	TTT	AAG	ATG	TGG	AAA	AGG	ATAC	TTT	GAT	AA	2940
T	AAA	AA	AA	AA	CAT	GG	AT	AA	AA	TTT	AA	AA	AA	AA	AA	AA	AA	TTT	TTT	3000
T	G	T	T	T	TT	T	T	T	T	T	T	T	T	T	T	T	T	T	3060	
T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	3120	
T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	3180	
T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	3181	
A																				

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AGCTCGGAAT	TCCGAGCTTG	GATCCCTCTAG	AGCGGGCCGCC	GACTAGTGAG	CTCGTCGACC	60
CGGGATT						68

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATTAATTCC	CGGGTCGACG	AGCTCACTAG	TCGGGGCCCG	CTCTAGAGGA	TCCAAGCTCG	60
GAATTCCG						68

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AGCGGATAAC	AATTTCACAC	AGGA	24
------------	------------	------	----

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGTAAACGCA	CGGCCAGT	18
------------	----------	----

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTGCCAGGCT AAAATTACGG

20

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATCACAGACC TTGATGCCAC

20

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCTGGTATTG TCTGTGATTG GGTC

24

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CATCAGGATT TGCTTGAGGG

20

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TATTGGTCAG GCAAGGGAAG

20

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

10025167 121901

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTGTTTGCTC CTCCATGAGC

20

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CAAGTAGAAAG GTCAGGGGTG

20

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATAAGTGTCA AGGAGGCAGC

20

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCAGACTGTT CCATGTGATG

20

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGTACCTGT TCTGGAGCC

20

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ACGTACCTGT TTGAAGCCAC

20

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGTAAGGACC GCCTAAATCG

20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GAAGTGAAAC AAAGTGGGGC

20

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TTATCCTCCC CATCAGTCAG

20

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TCGATTITAGG CGGTCCCTAC

20

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

10025167 121901

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGTGGCTTCA AACAGGTACG

20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GGGTAAGGAC CGCCTAAATC GAATG

25

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GAGCCCCAAA AGCATCAATG AGG

23

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 917 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Gly Leu Phe Arg Gly Phe Val Phe Leu Leu Val Cys Leu Leu
1 5 10 15
His Gln Ser Asn Thr Ser Phe Ile Lys Leu Asn Asn Asn Gly Phe Glu
20 25 30
Asp Ile Val Ile Val Ile Asp Pro Ser Val Pro Glu Asp Glu Lys Ile
35 40 45
Ile Glu Gln Ile Glu Asp Met Val Thr Thr Ala Ser Thr Tyr Leu Phe
50 55 60
Glu Ala Thr Glu Lys Arg Phe Phe Phe Lys Asn Val Ser Ile Leu Ile
65 70 75 80
Pro Glu Asn Trp Lys Glu Asn Pro Gln Tyr Lys Arg Pro Lys His Glu
85 90 95
Asn His Lys His Ala Asp Val Ile Val Ala Pro Pro Thr Leu Pro Gly
100 105 110
Arg Asp Glu Pro Tyr Thr Lys Gln Phe Thr Glu Cys Gly Glu Lys Gly
115 120 125
Glu Tyr Ile His Phe Thr Pro Asp Leu Leu Glu Lys Lys Gln Asn
130 135 140
Glu Tyr Gly Pro Pro Gly Lys Leu Phe Val His Glu Trp Ala His Leu
145 150 155 160

2020/07/26 10:52:00

Arg Trp Gly Val Phe Asp Glu Tyr Asn Glu Asp Gln Pro Phe Tyr Arg
165 170 175
Ala Lys Ser Lys Ile Glu Ala Thr Arg Cys Ser Ala Gly Ile Ser
180 185 190
Gly Arg Asn Arg Val Tyr Lys Cys Gln Gly Gly Ser Cys Leu Ser Arg
195 200 205
Ala Cys Arg Ile Asp Ser Thr Thr Lys Leu Tyr Gly Lys Asp Cys Gln
210 215 220
Phe Phe Pro Asp Lys Val Gln Thr Glu Lys Ala Ser Ile Met Phe Met
225 230 235 240
Gln Ser Ile Asp Ser Val Val Glu Phe Cys Asn Glu Lys Thr His Asn
245 250 255
Gln Glu Ala Pro Ser Leu Gln Asn Ile Lys Cys Asn Phe Arg Ser Thr
260 265 270
Trp Glu Val Ile Ser Asn Ser Glu Asp Phe Lys Asn Thr Ile Pro Met
275 280 285
Val Thr Pro Pro Pro Val Phe Ser Leu Leu Lys Ile Ser Gln
290 295 300
Arg Ile Val Cys Leu Val Leu Asp Lys Ser Gly Ser Met Gly Gly Lys
305 310 315 320
Asp Arg Leu Asn Arg Met Asn Gln Ala Ala Lys His Phe Leu Leu Gln
325 330 335
Thr Val Glu Asn Gly Ser Trp Val Gly Met Val His Phe Asp Ser Thr
340 345 350
Ala Thr Ile Val Asn Lys Leu Ile Gln Ile Lys Ser Ser Asp Glu Arg
355 360 365
Asn Thr Leu Met Ala Gly Leu Pro Thr Tyr Pro Leu Gly Gly Thr Ser
370 375 380
Ile Cys Ser Gly Ile Lys Tyr Ala Phe Gln Val Ile Gly Glu Leu His
385 390 395 400
Ser Gln Leu Asp Gly Ser Glu Val Leu Leu Leu Thr Asp Gly Glu Asp
405 410 415
Asn Thr Ala Ser Ser Cys Ile Asp Glu Val Lys Gln Ser Gly Ala Ile
420 425 430
Val His Phe Ile Ala Leu Gly Arg Ala Ala Asp Glu Ala Val Ile Glu
435 440 445
Met Ser Lys Ile Thr Gly Gly Ser His Phe Tyr Val Ser Asp Glu Ala
450 455 460
Gln Asn Asn Gly Leu Ile Asp Ala Phe Gly Ala Leu Thr Ser Gly Asn
465 470 475 480
Thr Asp Leu Ser Gln Lys Ser Leu Gln Leu Glu Ser Lys Gly Leu Thr
485 490 495
Leu Asn Ser Asn Ala Trp Met Asn Asp Thr Val Ile Ile Asp Ser Thr
500 505 510
Val Gly Lys Asp Thr Phe Phe Leu Ile Thr Trp Asn Ser Leu Pro Pro
515 520 525
Ser Ile Ser Leu Trp Asp Pro Ser Gly Thr Ile Met Glu Asn Phe Thr
530 535 540
Val Asp Ala Thr Ser Lys Met Ala Tyr Leu Ser Ile Pro Gly Thr Ala
545 550 555 560
Lys Val Gly Thr Trp Ala Tyr Asn Leu Gln Ala Lys Ala Asn Pro Glu
565 570 575
Thr Leu Thr Ile Thr Val Thr Ser Arg Ala Ala Asn Ser Ser Val Pro
580 585 590
Pro Ile Thr Val Asn Ala Lys Met Asn Lys Asp Val Asn Ser Phe Pro
595 600 605
Ser Pro Met Ile Val Tyr Ala Glu Ile Leu Gln Gly Tyr Val Pro Val
610 615 620
Leu Gly Ala Asn Val Thr Ala Phe Ile Glu Ser Gln Asn Gly His Thr
625 630 635 640
Glu Val Leu Glu Leu Leu Asp Asn Gly Ala Gly Ala Asp Ser Phe Lys
645 650 655
Asn Asp Gly Val Tyr Ser Arg Tyr Phe Thr Ala Tyr Thr Glu Asn Gly
660 665 670
Arg Tyr Ser Leu Lys Val Arg Ala His Gly Gly Ala Asn Thr Ala Arg
675 680 685

Leu Lys Leu Arg Pro Pro Leu Asn Arg Ala Ala Tyr Ile Pro Gly Trp
690 695 700
Val Val Asn Gly Glu Ile Glu Ala Asn Pro Pro Arg Pro Glu Ile Asp
705 710 715 720
Glu Asp Thr Gln Thr Thr Leu Glu Asp Phe Ser Arg Thr Ala Ser Gly
725 730 735
Gly Ala Phe Val Val Ser Gln Val Pro Ser Leu Pro Leu Pro Asp Gln
740 745 750
Tyr Pro Pro Ser Gln Ile Thr Asp Leu Asp Ala Thr Val His Glu Asp
755 760 765
Lys Ile Ile Leu Thr Trp Thr Ala Pro Gly Asp Asn Phe Asp Val Gly
770 775 780
Lys Val Gln Arg Tyr Ile Ile Arg Ile Ser Ala Ser Ile Leu Asp Leu
785 790 795 800
Arg Asp Ser Phe Asp Asp Ala Leu Gln Val Asn Thr Thr Asp Leu Ser
805 810 815
Pro Lys Glu Ala Asn Ser Lys Glu Ser Phe Ala Phe Lys Pro Glu Asn
820 825 830
Ile Ser Glu Glu Asn Ala Thr His Ile Phe Ile Ala Ile Lys Ser Ile
835 840 845
Asp Lys Ser Asn Leu Thr Ser Lys Val Ser Asn Ile Ala Gln Val Thr
850 855 860
Leu Phe Ile Pro Gln Ala Asn Pro Asp Asp Ile Asp Pro Thr Pro Thr
865 870 875 880
Pro Thr Pro Thr Pro Asp Lys Ser His Asn Ser Gly Val Asn Ile Ser
885 890 895
Thr Leu Val Leu Ser Val Ile Gly Ser Val Val Ile Val Asn Phe Ile
900 905 910
Leu Ser Thr Thr Ile
915

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Ala Asn Ser Ser Val Pro Pro Ile Thr Val Asn Ala Lys Met Asn Lys
1 5 10 15
Asp Val Asn Ser Phe
20

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Asp Asn Gly Ala Gly Ala Asp Ser Phe Lys Asn Asp Gly Val Tyr Ser
1 5 10 15
Arg Tyr Phe Thr Ala Tyr Thr Glu Asn Gly Arg Tyr Ser Leu Lys
20 25 30

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Val Arg Ala His Gly Gly Ala Asn Thr Ala Arg Leu Lys Leu Arg Pro
1 5 10 15
Pro Leu Asn Arg Ala Ala Tyr Ile
20

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Ser Leu Pro Leu Pro Asp Gln Tyr Pro Pro Ser Gln Ile Thr Asp Leu
1 5 10 15
Asp Ala Thr Val His Glu Asp Lys Ile Ile Leu Thr Trp Thr Ala Pro
20 25 30
Gly Asp Asn Phe Asp Val Gly Lys
35 40

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Tyr Asn Glu Asp Gln Pro Phe Tyr Arg Ala Lys Ser Lys Lys Ile Glu
1 5 10 15
Ala Thr Arg Cys
20

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Leu Ser Arg Ala Cys Arg Ile Asp Ser Thr Thr Lys Leu Tyr Gly Lys
1 5 10 15
Asp Cys Gln Phe Phe Pro Asp Lys
20

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Lys Ser Ser Asp Glu Arg Asn Thr Leu Met Ala Gly Leu Pro Thr Tyr
1 5 10 15
Pro Leu Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Glu Ile Asp Glu Asp Thr Gln Thr Thr Leu Glu Asp Phe Ser Arg
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

1002517755001

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met His Thr Glu His
1 5 10 15
His His His His His
20